DOCUMENT RESUME

ED 109 198 TM 004 652

AUTHOR Besel, Ronald
TITLE An Initial Facet Analysis of the FYCSP Word Attack Test.
INSTITUTION Southwest Regional Laboratory for Educational Research and Development, Los Alamitos, Calif.
REPORT NO TN-5-72-46
PUB DATE 18 Aug 72
NOTE 17p.

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Guessing (Tests); *Item Analysis; *Multiple Choice Tests; Predictor Variables; *Response Style (Tests); Stimuli; Testing; *Tests; *Word Recognition
IDENTIFIERS Distractors (Tests); *Facet Analysis; FYCSP Word Attack Test

ABSTRACT

The concepts "facet analysis," "facet design," and "facet structure" are defined. The FYCSP (First Year Communication Skills Program) Word Attack Test is analyzed in terms of two related facet structures. Stepwise linear regression is used to predict distractor attractiveness. Hypotheses suggested by Guttman relating distractor attractiveness to "degree of similarity to the correct answer" are formulated and tested. Potential applications of facet analysis are discussed. (Author/DEP)

SWRL

SOUTHWEST REGIONAL LABORATORY
TECHNICAL NOTE

DATE: TN 5-72-46

NO: August 18, 1972

TITLE: AN INITIAL FACET ANALYSIS OF THE FYCSP WORD ATTACK TEST

AUTHOR: Ronald Besel

ABSTRACT

The concepts "facet analysis," "facet design," and "facet structure" are defined. The FYCSP Word Attack Test is analyzed in terms of two related facet structures. Stepwise linear regression is used to predict distractor attractiveness. Hypotheses suggested by Guttman relating distractor attractiveness to "degree of similarity to the correct answer" are formulated and tested. Potential applications of facet analysis are discussed.

4665 Lampson Ave. Los Alamitos, Ca.

AN INITIAL FACET ANALYSIS OF THE
FYCSP WORD ATTACK TEST

Facet analysis is a collection of procedures for describing the content of multiple-choice test items and interpreting the observed item response patterns. The content of a test item is assumed to have two aspects: the stimulus and the response options. For the FYCSP word attack test, the stimulus is a spoken word which can be constructed from word elements using one or more phonetic rules. For a reading comprehension test, the stimulus may be a selected reading text (Schlesinger and Weiser, 1970). A facet is defined to be a characteristic on which the stimulus and an option can be evaluated and compared.

Facet analysis involves identifying characteristics (facets) and hypothesizing which are significant in explaining response patterns. Response patterns can be described quantitatively several ways. A similarity measure, s_i, for each facet is defined:

$$s_i = \begin{cases} 1 & \text{if the option and the stimulus are identical with respect to facet (i)} \\ 0 & \text{if the option and the stimulus are different} \end{cases}$$

If more than one facet is used to describe a distractor a similarity vector is defined:

$$S = (s_1, s_2, \ldots s_n) \qquad (1)$$

The proportion of respondents that selected a distractor is referred to as the distractor attractiveness.

Data Source

The data available for analysis were the FYCSP criterion exercises from a sample of 10 classrooms selected from the 1970-71 Quality Assurance project. Data from 362 students were available from the first two units with the sample size shrinking to 98 students by Unit 10. The initial use of this data was for developing procedures to handle error conditions in IMS input data (Hooper and McManus, 1971). To meet the needs of that study, criterion exercise data were scored using procedures simulating the output of optical scanner interpretation of student responses. Thus, some responses were scored as multiple-marks in instances where a human could easily distinguish that the student meant to cross-off or erase one response. However, this was not viewed as consequential for the purposes of this preliminary analysis.

The sample of test items consists of the 50 word attack items (five-per-unit) of the 1970-71 FYCSP criterion exercises. Each item had three response options resulting in a total sample size of 100 distractors.

Possible Facet Structures

All of the stimuli for the word attack tests are composed of three grapheme units--an initial consonant sound, a vowel sound and a final consonant sound. In word attack instruction, however, each word is treated as an initial consonant sound and a final vowel-consonant sound. Thus, both a three facet (C-V-C) and a two facet (C-VC) structure appear plausible.

The three facet structure results in seven classes of distractors, all of which are represented in the sample of 100 distractors but with a very uneven distribution (see Table 1). The three classes of distractors for the two facet structure are formed by collapsing the three facet classes:

$$(0,1,1) \Rightarrow (0,1)$$

$$\begin{bmatrix}(1,1,0)\\(1,0,1)\\(1,0,0)\end{bmatrix} \Rightarrow (1,0)$$

$$\begin{bmatrix}(0,1,0)\\(0,0,1)\\(0,0,0)\end{bmatrix} \Rightarrow (0,0)$$

For example, if the word "sat" is sounded out by the teacher as a stimulus, possible distractors would be classified as follows:

Distractor	Two Facet Structure	Three Facet Structure
Mat	(0,1)	(0,1,1)
sit	(1,0)	(1,0,1)
hit	(0,0)	(0,0,1)
I	(0,0)	(0,0,0)
see	(1,0)	(1,0,0)
has	(0,0)	(0,1,0)
sap	(1,0)	(1,1,0)

Guttman's Hypothesis

Guttman has suggested the following hypothesis,

"...the degree of attraction of a distractor increases monotonely with its 'degree of similarity' to the correct answer" (Guttman and Schlesinger, 1967).

Let $A(S)$ represent the hypothesized relationship between degree of similarity and attractiveness, the following partial orders are then predicted:

$$A(1,1,0) > \begin{Bmatrix} A(1,0,0) \\ A(0,1,0) \end{Bmatrix} > A(0,0,0) \tag{2}$$

$$A(1,0,1) > \begin{Bmatrix} A(1,0,0) \\ A(0,0,1) \end{Bmatrix} > A(0,0,0) \tag{3}$$

$$A(0,1,1) > \begin{Bmatrix} A(0,1,0) \\ A(0,0,1) \end{Bmatrix} > A(0,0,0) \tag{4}$$

Guttman actually proposed a somewhat stronger hypothesis than these partial orders. Define the level (L) of a distractor to be the number of s_i values for an item equal to one,

$$L = s_1 + s_2 + s_3 \ldots + s_n \tag{5}$$

then the attractiveness of a distractor is predicted to increase monotonely with respect to level.

Predicted Effects of Other Variables

It is expected that other variables, in addition to distractor-stimulus similarity, may affect attractiveness. It generally is anticipated that the attractiveness of distractors will decrease as a student progresses through a program due to learning. It is expected that this rate of decrease will not be constant but be greatest for the initial units.

It is expected that, if the student does not know the correct answer, he will have a tendency to pick the first option. As a result of learning, this tendency should be less in evidence in later unit tests.

TABLE 1

Frequency of Occurrence of Distractor Types as a Function of Unit and Position

(Three-Facet Structure)

UNIT	Type of Distractor Level 2 (1,1,0)	(1,0,1)	(0,1,1)	Level 1 (1,0,0)	(0,1,0)	(0,0,1)	Level 0 (0,0,0)
1	0	0	3 2-0-1	5 2-2-1	1 0-0-1	0	1 0-1-0
2	1 1-0-0	2 1-0-1	4 0-3-1	2 0-1-1	1 1-0-0	0	0
3	2 1-0-1	1 0-1-0	5 2-1-2	1 0-1-0	0	1 0-1-0	0
4	0	2 1-0-1	4 0-2-2	2 1-0-1	1 0-0-1	1 1-0-0	0
5	0	2 0-1-1	5 1-4-0	1 0-0-1	0	2 2-0-0	0
6	0	1 0-0-1	5 1-2-2	2 0-0-2	1 1-0-0	1 0-1-0	0
7	1 1-0-0	2 2-0-0	5 1-1-3	2 0-0-2	0	0	0
8	2 2-0-0	0	4 1-1-4	3 0-1-2	0	0	1 1-0-0
9	3 1-1-1	0	4 2-1-1	1 0-0-1	1 0-0-1	1 0-1-0	0
10	1 0-0-1	1 1-0-0	5 2-2-1	2 0-2-0	1 1-0-0	0	0
Totals	10 6-1-3	11 5-2-4	44 12-17-15	21 3-7-11	6 3-0-3	6 3-3-0	2 1-1-0

Notes: Total for a distractor type for a unit is given in the upper number. Positional breakdown, left-center-right, is given below if the cell was not empty.

Variables are defined in such a way that their first order effects in predicting attractiveness will have predicted positive Beta weights in a linear regression equation.

U = 10 - unit number for test containing the item.

$$P_D = \begin{cases} 1 & \text{if the distractor is the first option.} \\ 0 & \text{if the distractor is the second option.} \\ -1 & \text{if the distractor is the third option.} \end{cases}$$

The similarity of the two distractors for a item is expected to have a secondary effect on the attractiveness of each distractor. It is hypothesized that inter-distractor similarity reduces the attractiveness of each distractor. Thus, the most attractive distractor is predicted to be one which is both maximally similar to the correct response and different from its paired distractor. A distractor similarity variable (DS) is defined to be equal to the number of facets for which the two distractors are identical.

Regression Analysis

Table 1 gives the distribution of distractor type by unit and by position for the three-facet structure. The column totals indicate that the overall occurrence of distractor types is quite unbalanced. Furthermore, the occurrence of some of the distractor types does not appear to be uniform across either unit or position. Comparison of observed mean attractiveness could be misleading due to the high frequency of empty cells and a confounding of effects. The hypothesized effects of similarity, unit and position were felt to be best tested by comparing the predicted attractiveness of classes of distractors.

Stepwise linear regression was employed to generate predictive equations of increasing complexity. As displayed in Tables 3 and 4, a linear model with main effects only was first estimated. All of the Beta weights were positive as predicted. This indicates that, if only the main effects are considered, each of the variables effects the attractiveness of a distractor in the predicted direction.

Adding the first order interaction effects significantly increases the variance accounted for by both regression equations. As predicted, the Beta weights for the interaction of unit and position (UxP_D) are positive. This is the expected result if the position effect is most in evidence for the tests early in the program.

The interaction of a similarity variable, for a facet, with the unit variable is predicted to have a positive Beta weight if it is expected that the student population will evidence a reduced tendency to pick distractors, with that similarity, on later unit tests. Thus, if students learn to reject every class of distractors, positive Beta weights for all (Uxs_i) terms are predicted. If, however, students tend to learn to reject some classes of distractors and not others, then some of the Beta weights may be zero or negative. It is seen that the Beta weights for (Uxs_1)--interaction of unit with initial consonant sound--is negative. Possible implications of this result will be discussed later.

The distractor similarity (DS) variable and the second order unit effect (U^2) were added last to the three facet model. The variance accounted for was significantly increased with most of the increase due to (U^2). The sign of the U^2 Beta weight was positive as predicted.

Examination of the type of distractor similarities occurring revealed that there is probably a confounding of distractor similarity with the s_i variables. When the two distractors in an item exhibited a similarity in one or more facets, with but a single exception, they shared that similarity with the correct option. If the two distractors had two identical facets, they were always the vowel, final-consonant pair. There was only one occurrence of a distractor similarity occurring in the initial consonant sound. Due to the apparent confounding, interpretation of the distractor similarity Beta weight was not attempted and the second order effects were deleted from the two-facet analysis.

Discussion of Results

Table 2 contains the predicted attractiveness of each class of distractors, for each unit, for the three-facet model with the position and distractor similarity variables set to zero. Figures 1 and 2 compare the observed mean attractiveness of the two most frequently occurring classes with the predicted values.

The partial order hypotheses of the relative attractiveness of distractors are supported by the Table 2 values. Of the 90 partial order predictions, only seven differ from the values predicted by the regression equation. Of the 30 additional order predictions generated by the "level-of-distractor" hypotheses, 4 differ from the regression predictions. These 4, however, all occurred in the comparison of the two most frequently occurring classes (Figures 1 and 2) and thus constitute stronger evidence of the weakness of the "level-of-distractor" hypotheses.

TABLE 2

Predicted Attractiveness of Distractors and Frequency of their Occurrence for the FYCSP Word Attack Test.

Unit	Type of Distractor						
	Level 2			Level 1			Level 0
	(1,1,0)	(1,0,1)	(0,1,1)	(1,0,0)	(0,1,0)	(0,0,1)	(0,0,0)
1	.118 (0)	.141 (0)	.184 (3)	.068 (5)	.103 (1)	.094 (0)	.034 (1)
2	.097 (1)	.130 (2)	.150 (4)	.056 (2)	.077 (1)	.071 (0)	.018 (0)
3	.077 (2)	.113 (1)	.120 (5)	.047 (1)	.055 (0)	.051 (1)	.005 (0)
4	.061 (0)	.100 (2)	.094 (4)	.042 (2)	.036 (1)	.035 (1)	-.033 (0)
5	.050 (0)	.091 (2)	.072 (5)	.040 (1)	.021 (0)	.022 (2)	-.008 (0)
6	.042 (0)	.085 (1)	.052 (5)	.042 (2)	.010 (1)	.013 (1)	-.010 (0)
7	.037 (1)	.084 (2)	.037 (5)	.048 (2)	.002 (0)	.008 (0)	-.007 (0)
8	.036 (2)	.085 (0)	.024 (4)	.057 (3)	-.003 (0)	.006 (0)	-.003 (1)
9	.039 (3)	.090 (0)	.017 (4)	.070 (1)	-.003 (1)	.010 (1)	.009 (0)
10	.046 (1)	.100 (1)	.013 (5)	.086 (2)	.000 (1)	.014 (0)	.022 (0)
Total Number of Distractors	10	11	44	21	6	6	2

Notes: (1) The frequency of occurrence of each type of distractor by unit indicated by the numbers in parentheses.

(2) The Beta weights for the full 3-facet model (third column of Table 3) were used with distractor similarity and position set to zero.

Figure 1. Attractiveness (A) as a function of program unit for distractors with correct initial sound, type (1,0,0)

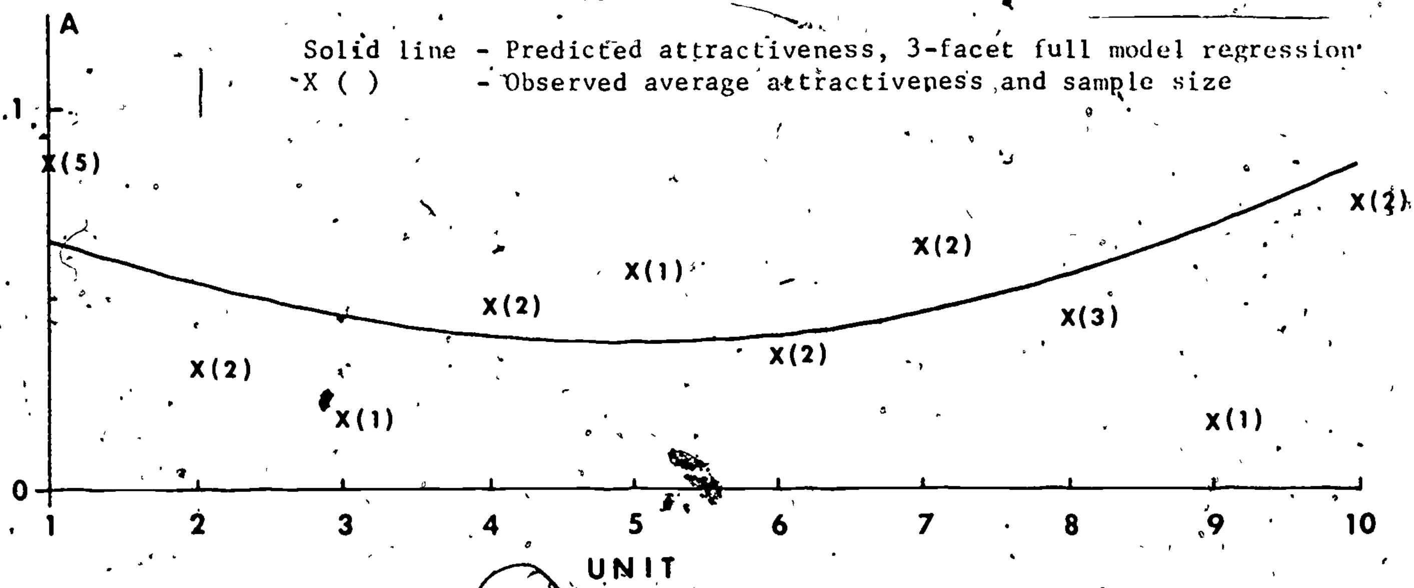
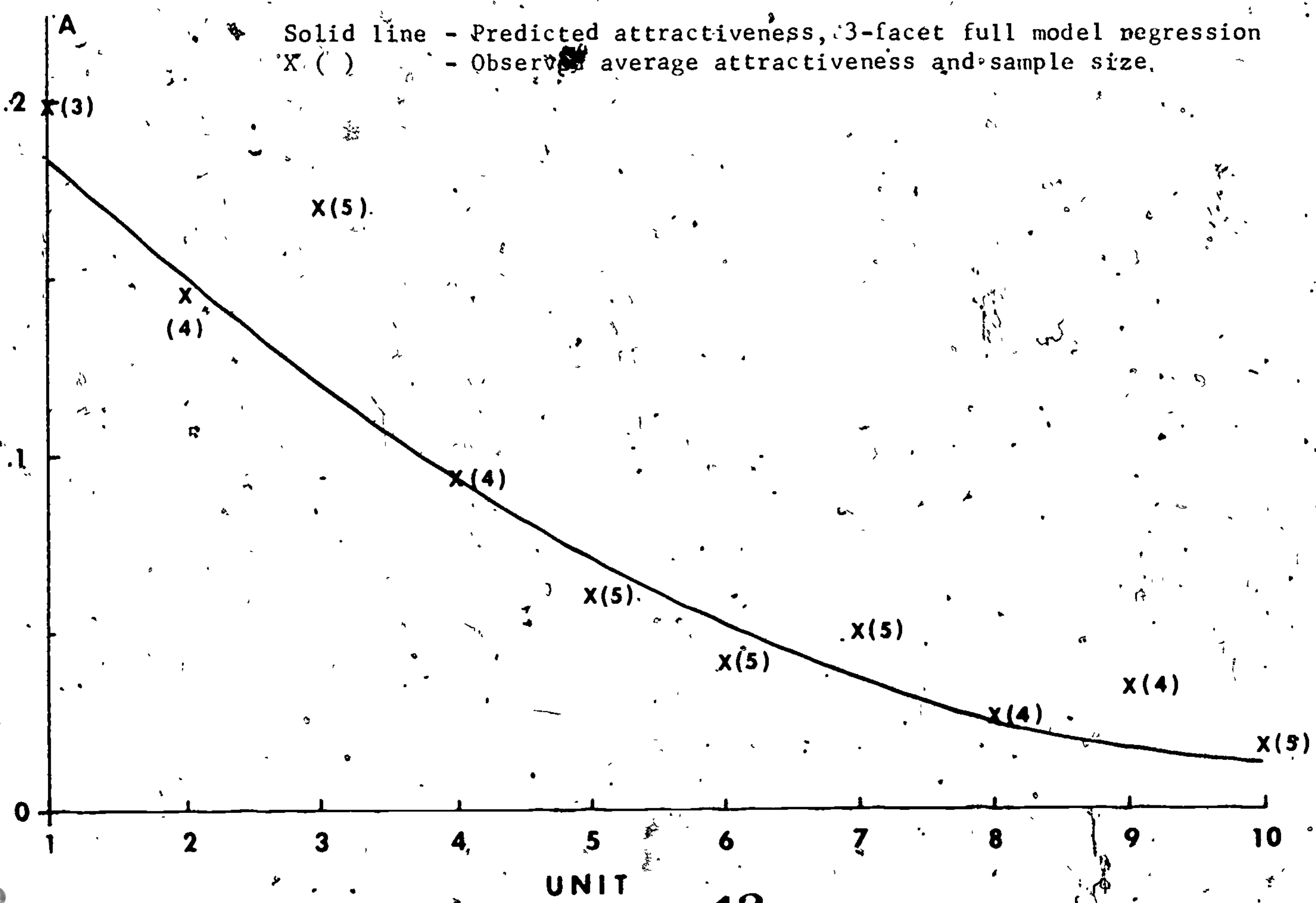


Figure 2. Attractiveness (A) as a function of program unit for distractors with correct vowel-consonant final sound, type (0,1,1)



It seems likely that the "level-of-distractor" hypotheses will not be supported in any instance where one facet influences attractiveness to a substantially lesser degree than the other facets. For the word attack test, the relative effects of the facets seemed to vary over time.

Differential learning is a potential interpretation of the noticable differences in slope between the regression equations for the two distractor types plotted in Figures 1 and 2. If such is the case, inclusion of more practice in decoding the medial vowel and final consonant sounds of a word may be a desirable program modification. It is possible, however, that the observed changes in distractor attractiveness across units is due to a tendency to speed up test administration in the later units.

Facet Design

Systematic rules for constructing sets of distractors according to an a-priori choice of facets constitutes a facet design. Tests based on a facet design have the following advantages over the usual multiple choice tests. (Guttman and Schlesinger, 1967).

1. Successful prediction of relative empirical difficulties of distractors.
2. Reduction of variation in test results due to undesired factors.
3. Possibility of differential scoring of subjects on the types of wrong answers to which they are attracted.

The first two advantages relate to the potential for improving the reliability and validity without increasing test length. The third refers

to the possibility of gaining diagnostic information from an analysis of the errors a student makes.

A potential area for both theoretical and empirical research is the determination of which types of facet designs are optimal for improving test validity. An unanswered question is whether improving test validity increases or decreases the amount of diagnostic information that can be gleaned from error patterns.

Potential Applications of Facet Analysis to IMS

Facet analysis may yield benefits for each of the areas currently included in the IMS effort. Construction of tests according to a facet design would permit a systematic investigation of this potential. If the student population as a whole exhibits differential learning in rejecting classes of distractors, there are implications for program revision. If further analysis reveals that individual students tend to select distractors from some class or classes in a predictable fashion, individualized diagnosis and prescription may be indicated.

A facet design leads to the construction of test items on the basis of an a-priori definition of test content. This should be a reasonable starting point for developing computer generated tests and computer generated drill-and-practice exercises.

TABLE 3

Linear Regression Models for the Three Facet Prediction of Distractor Attractiveness (Beta Weights)

	Variable	Main Effects Only	First Order Interactions	Second Order Effects
	Constant Term	-.0425	.0074	.0219
Main Effects	Position (P_D)	.0104	-.0029	-.0056
	Similarity, Facet 1 (s_1)	.0505	.0595	.0645
	Similarity, Facet 2 (s_2)	.0317	-.0153	-.0214
	Similarity, Facet 3 (s_3)	.0468	-.0193	-.0077
	Unit (U)	.0099	.0024	-.0151
First Order Interactions	$U \times P_D$		.0039	.0039
	$U \times s_1$		-.0039	-.0033
	$U \times s_2$		.0085	.0101
	$U \times s_3$		.0069	.0076
	$s_1 \times s_2$		-.0126	-.0194
	$s_1 \times s_3$		.0359	.0208
	$s_2 \times s_3$		.0310	.0202
Second Order Effects	Unit Squared (U^2)			.0018
	Distractor Similarity (DS)			.0114
	Multiple Correlation:	.633	.781	.816
	Variance Accounted For:	.400	.610	.666

TABLE 4

Linear Regression Models for the Two Facet Prediction of Distractor Attractiveness (Beta Weights)

Variable			Main Effects Only	First Order Interactions
Constant Term			-.0120	.0147
Main Effects	Position	(P)	.0141	.0000
	Similarity, Facet 1	(s_1)	.0417	.0492
	Similarity, Facet 2	(s_2)	.0509	.0176
	Unit	(U)	.0096	.0042
First Order Interactions	U x P			.0034
	U x s_1			-.0021
	U x s_2			.0152
Multiple Correlation:			.592	.748
Variance Accounted For:			.351	.559

REFERENCES

Guttman, L. and Schlesinger, I.M., "Systematic Construction of Distractors for Ability and Achievement Test Items," Educational and Psychological Measurement, 1967, Vol. 2, 569-580.

Hooper, J. and McManus, J.F., "IMS Error Conditions, Their Identification and Resolution," SWRL TN 5-71-11, March 1971.

Schlesinger, I.M. and Weiser, Z., "A Facet Design for Tests of Reading Comprehension," Reading Research Quarterly, 1970, Vol. 5, 566-580.